

## **PRODUCTION OF 2'-DEOXYNUCLEOSIDES AND 2'-DEOXYNUCLEOSIDE PRECURSORS FROM 2-DEHYDRO-3-DEOXY-D-GLUCONATE**

This invention relates to a process for preparing 2'-deoxynucleoside compounds or 2'-deoxynucleoside precursors using 2-dehydro-3-deoxy-D-gluconic acid (usually abbreviated as KDG) or its salts as a starting material. A variety of 2'-deoxynucleosides and their analogues are used as a starting material for synthesis or drug formulation in production of an antiviral, anticancer or antisense agent.

Specifically, the invention relates to a method in which KDG or a derivative of KDG is subjected to a decarboxylation step to remove the original carboxy group of KDG. In a preferred embodiment, the KDG used in the method according to the invention is enzymatically produced from D-gluconate or D-glucosamine.

2'-deoxynucleosides and 2'-deoxynucleoside precursors including 2-deoxy-D-ribose are used as starting material for synthesis or drug formulation, for instance, in production of antiviral and anticancer agent. 2'-deoxynucleosides or derivatives thereof and 2'-deoxynucleoside precursors are also used as reagents for research, diagnosis and synthesis of therapeutic antisense molecules.

In one method of the prior art, deoxynucleosides are generated from biological materials such as testis (WO 99/49074) or yeast or fish sperm by enzymatic cleavage of DNA. This method, however, involves several disadvantages, in particular regarding difficulties of obtaining the starting material in sufficient quantity and quality.

The main production process of 2-deoxy-D-ribose currently consists in chemical hydrolysis of DNA. In this case, the deoxyribosyl moiety originates in ribonucleotide reductase activity. No synthesis of 2-deoxy-D-ribose from KDG has been yet described.

In most living cells, deoxyribonucleosides result from a "salvage pathway" of the

nucleotide metabolism. The deoxyribose moiety of deoxyribonucleosides is obtained through the reduction of a ribosyl moiety into di- or triphosphate ribonucleotides catalyzed by ribonucleotide reductases. However, the deoxyribose moiety is not recycled, but is degraded into D-glyceraldehyde-3-phosphate and acetaldehyde following the reactions of central metabolism:

- deoxynucleoside is cleaved into deoxyribose-1-phosphate and nucleobase through phosphorolysis mediated by products of the genes encoding thymidine phosphorylase (deoA), purine-nucleoside phosphorylase (deoD), uridine phosphorylase (udp) or xanthosine phosphorylase (xapA).
- deoxyribose-1-phosphate is converted into deoxyribose-5-phosphate through a reaction catalyzed by deoxyribose phosphate mutase (deoB),
- which is further degraded to D-glyceraldehyde-3-phosphate and acetaldehyde through a reaction catalyzed by deoxyribose-5-phosphate aldolase (deoC).

It has been shown that the deo enzymes also catalyze in vitro the reverse anabolic reactions: Deoxyribose-5-phosphate is obtained in vitro in the presence of purified *Escherichia coli* or *Lactobacillus plantarum* deoxyribose aldolase starting from acetaldehyde and D-glyceraldehyde-3-phosphate (Rosen et al., J. Biol. Chem., 240, (1964), 1517-1524; Pricer, J. Biol. Chem., 235, (1960), 1292-1298). Deoxyribose can also be obtained with acetaldehyde and glyceraldehyde as enzyme substrates, but only with a very low yield (Barbas, J. Am. Chem. Soc. 112 (1990), 2013-2014).

The patent application WO 01/14566 describes the enzymatic synthesis of deoxynucleosides starting from deoxyribose-1-phosphate through the combined activities of three enzymes of the deo operon, i.e. deoxyribose aldolase, deoxyribomutase and phosphorylase (thymidine or purine nucleoside phosphorylase) in a one-pot reaction, using as starting substrates glyceraldehyde-3-phosphate, acetaldehyde and a nucleobase. D-glyceraldehyde-3-phosphate can be obtained from fructose-1,6-bisphosphate by an enzymatic process.

The patent application EP 1179598 describes the use of phosphorylase to catalyze the enzymatic production of deoxynucleosides starting from deoxyribose-1-phosphate and nucleobase. The yield of deoxynucleoside synthesis is improved by

precipitation of phosphate.

However, methods using enzymes of the deo operon working in the reverse direction compared to their biological function show low yields, which indicates serious drawbacks for their use.

In view of the above-described ineffectiveness of the currently applied processes for producing deoxynucleosides and deoxynucleoside precursors, it is an object of the present invention to provide means and methods for the biosynthetic production of deoxynucleosides and deoxynucleoside precursors starting from cheap and commercially available compounds without being dependent on unreliable natural sources.

In particular, there is a need for alternative methods for the production of deoxynucleosides and deoxynucleoside precursors which allow efficient and economical synthesis of deoxyribonucleosides, by means of which the drawbacks of prior art processes are eliminated.

The present invention relates to a method for producing 2'-deoxynucleosides and precursors thereof starting from 2-dehydro-3-deoxy-D-gluconic acid (KDG) or its salts and comprising a decarboxylation step.

In particular, this method is useful for producing 2-deoxy-D-ribose (DRI) as well as synthetically versatile enamine derivatives of DRI as 2'-deoxynucleoside precursors.

The decarboxylation step takes place by reacting either KDG or its salts directly, or a derivative of KDG, usually to cleave the C1-C2 bond of the KDG.

In one embodiment of the invention, KDG or one of its salts undergoes (oxidative) decarboxylation leading to 2-deoxy-D-ribonic acid (DRN) or its salts, itself being further converted into 2-deoxy-D-ribose (DRI) or 2-deoxy-D-ribitol (DRL).

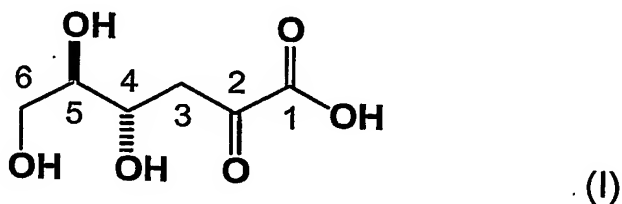
In another embodiment of the invention, decarboxylation takes place by reacting KDG or its salts with an amine, leading to an enamine derivative. This high energy enamine derivative can be further converted into DRI by hydrolysis.

In another embodiment of the invention, (oxidative) decarboxylation is carried out on 3-deoxy-D-gluconic acid (DGN) or its salts and/or 3-deoxy-D-mannonic acid (DMN) or its salts as derivatives of KDG, leading to DRI. Production of a mixture of DGN and DMN takes place by reduction of KDG. The decarboxylation is preferably carried out via reaction with hydrogen peroxide.

In another embodiment of the invention, (oxidative) decarboxylation is carried out on 3-deoxy-D-glucosaminic acid (DGM) or its salts and/or 3-deoxy-D-mannosaminic acid (DMM) or its salts, leading to DRI. Production of a mixture of DGM and DMM takes place from KDG by reductive amination.

Another aspect of the invention is a convenient and cost-effective method for preparing KDG or its salts to be used in the above methods. This method starts either from D-gluconate or from D-glucosamine through the use of recombinant enzymes. The invention provides a novel nucleotide sequence encoding a polypeptide having D-gluconate dehydratase activity and a nucleotide sequence encoding a polypeptide having D-glucosamine deaminase activity.

The starting material used for the method of the present invention is KDG, represented by formula (I) below or one of its salts, or a protected derivative thereof wherein one or more of the hydroxyl groups at positions 4, 5 and/or 6 are protected by a protection group known in the art.



The term "2'-deoxynucleoside" as used herein relates to 2'-deoxyribonucleosides which are N-glycosides, and wherein the basic N-atom of the nucleobase or nucleobase analog is bound to the anomeric carbon atom of 2-deoxy-D-ribose, or one of its derivatives. Examples of a suitable nucleobase are adenine, cytosine, guanine, thymine, uracil, 2,6-diaminopurine, and hypoxanthine. Examples of nucleobase analogs are 5-azacytosine, 2-chloro-adenine, 5-iodo-cytosine, 8-azaguanine, 5-iodo-uracil, 5-bromo-uracil, 5-fluoro-uracil, 5-ethyl-uracil and 5-trifluoromethyl-uracil.

The term "2'-deoxynucleoside precursors" as used herein, relates to compounds which can be easily converted into 2'-deoxynucleosides by applying methods known in the prior art. Preferred 2'-deoxynucleoside precursors are 2-deoxy-D-ribose (DRI) or carbohydrate compounds which can be converted into the 2-deoxy-D-ribosyl moiety of 2'-deoxynucleosides, for instance, those established in the prior art 1-phospho-2-deoxy-D-ribose, 5-phospho-2-deoxy-D-ribose and those established by the present invention 2-deoxy-D-ribitol, 2-deoxy-D-ribonic acid, 2-deoxy-D-ribono-1,4-lactone, 1-N-morpholino-3,4,5-trihydroxy-pentene-1, and their derivatives.

The method of the invention encompasses methods wherein the decarboxylation step is directly carried out on KDG or its salts or on compounds derived from KDG. Preferred KDG derivatives are 3-deoxy-D-gluconic acid, 3-deoxy-D-mannonic acid, 3-deoxy-D-glucosaminic acid and 3-deoxy-D-mannosaminic acid and their respective salts.

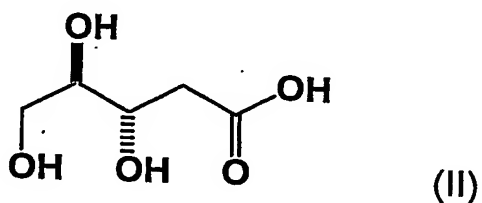
Furthermore, KDG and its salts or protected forms of these wherein one or more of the hydroxyl groups at the positions 4,5 and /or 6 are replaced by protecting groups known for that purpose in the art are also suitable starting materials for the decarboxylation reaction of the present invention. Unless noted otherwise, any reference to KDG in the following specification embraces protected forms of KDG, just as reference to KDG derivatives is intended to embrace protected forms of these derivatives. Similarly, any reference to the products obtained in the methods of the invention is intended to encompass protected forms of these products. Preferred

protection groups for the purpose of the invention are those which replace the respective hydroxyl groups by acetate ester, benzoate ester, allyl ether, benzyl ether, trityl ether, ter-butyldimethylsilyl (TBDMS) ether, isopropylidene or a benzylidene acetal.

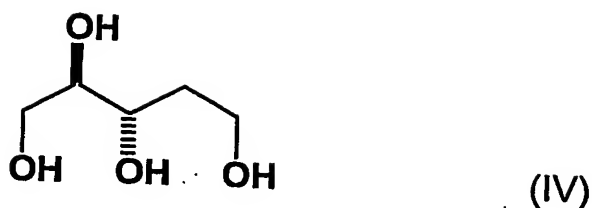
It should be understood that, depending on suitable reaction conditions for the embodiments of the invention, the carboxylic groups contained in the organic acids used as reactants or obtained as products can be in a protonated form or in their salt form, or may be present in equilibrium. Exemplary salts of these acids are those which have metal or ammonium ions as counterions, particularly alkali metal ions such as sodium and/or potassium.

Most of the carbohydrate compounds and their derivatives described in the present invention exist under several cyclic form but for simplicity reasons have been represented by open chain formulas. It is understood that the present invention encompasses all these isomeric or tautomeric forms.

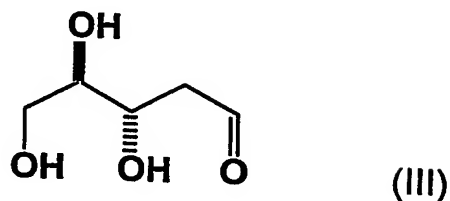
In a first embodiment of the invention, KDG or its salts is reacted with hydrogen peroxide and undergoes (oxidative) decarboxylation to 2-deoxy-D-ribonic acid (DRN), a compound of formula (II) or its salts.



The product may be further converted into or 2-deoxy-D-ribitol (DRL), represented by formula (IV)



or 2-deoxy-D-ribose (DRI), represented by formula (III)



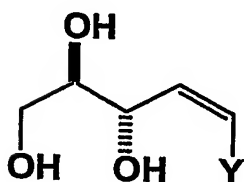
DRN, DRL and particularly DRI are among preferred 2'-deoxynucleoside precursors for the purpose of the present invention. Conversion of DRN to DRI may proceed directly or via DRL as an intermediate.

Preferably, the preparation of DRN is carried out by oxidative decarboxylation of sodium or potassium 2-dehydro-3-deoxy-D-gluconate in aqueous solution with hydrogen peroxide at room temperature as described in example 5. A general method for the preparation of aldonic acids by oxidative decarboxylation of 2-ketoaldonic acids is described in patent EP 1 038 860 A1.

Preferably, the preparation of DRL is carried out by hydrogenation of 2-deoxy-D-ribonolactone in aqueous solution with Rhodium catalyst on carbon at a temperature of 130°C under a pressure of 80 bars as described in example 6. 2-Deoxy-D-ribonolactone can be easily prepared by converting a 2-deoxy-D-ribonate (DRN salt) into 2-deoxy-D-ribonic acid, which is in equilibrium with its lactonic form in aqueous solutions (Han, Tetrahedron. 1993. 49, 349-362; Han, Tetrahedron Asymmetry. 1994. 5, 2535-62).

Preferably the preparation of 2-deoxy-D-ribose (DRI) is carried out by oxidization of 2-deoxy-D-ribitol (DLR), e.g. with chromium oxide in pyridine.

In another embodiment of the invention, decarboxylation takes place by reacting (KDG) or its salts with an amino group-containing reagent Y-H leading to a compound of formula (V).



(V)

or its respective trans isomer or a protected form thereof, as a 2'-deoxynucleoside precursor. Y-H represents an amine with the hydrogen atom H bound to the nitrogen of the amino group.

In a preferred embodiment of the invention, the amino group-containing reagent represented by Y-H is a linear or cyclic secondary amine; a primary amine that possess a  $\beta$ -carbonyl group, preferably 3-amino-2-indolinone which was found to be effective for the decarboxylation of  $\alpha$ -keto acids (Hanson, J. Chem. Education, 1987, 591-595). In each of these cases, -Y in formula (V) represents the respective nitrogen containing residue derived from these amino-group containing reagent.

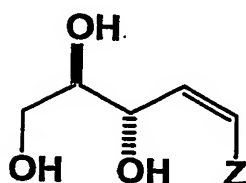
Preferably, the compound of formula (V) represents an enamine produced via reaction of a linear or cyclic secondary amine as Y-H.

Preferred cyclic secondary amines are morpholine, pyrrolidine, piperidine, or N-methyl piperazine; preferred non-cyclic amines are those of the formula  $R_1\text{-NH-R}_2$ , wherein  $R_1$  and  $R_2$  independently represent a linear or branched alkyl group of 1-8, preferably 1 to 4 carbon atoms. Particularly preferred as a non-cyclic amine is diethylamine.

Particularly preferred as a cyclic amine is morpholine.



The compound of formula (V) or its trans isomer or a protected form thereof can be further reacted with Z-H, wherein H represents a hydrogen atom and Z represents a leaving group, to produce a compound of formula (VI)



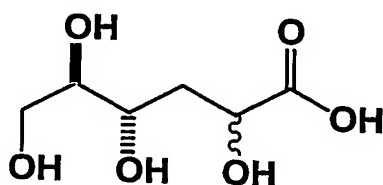
(VI)

or its respective trans isomer or a protected form thereof, as a 2'-deoxynucleoside precursor. Z-H is preferably water, in which case the compound of formula (VI) is DRI or a protected form thereof (keto-enol-tautomerism).

Preferably, the preparation of the compound of formula (V) is carried out by reacting KDG in benzene with the amine, e.g. morpholine under reflux using the method described in example 7, leading to 1-N-morpholino-3,4,5-trihydroxy-pentene-1. Acid catalysed hydrolysis with water yields 2-deoxy-D-ribose (DRI)

A general route to aldehydes via enamines from  $\alpha$ -oxocarboxylic acids carrying  $\beta$ -hydrogens is described by Stamos (Tetrahedron Lett. 23 (1982), 459-462). Other methods for the preparation and hydrolysis of enamines have been described elsewhere (Stork, J. Am. Chem. Soc. 85 (1963), 207-222; Stamhuis, J. Org. Chem. 30 (1965), 2156-2160).

In another embodiment of the invention, KDG or its salt is converted to 3-deoxy-D-gluconic acid (DGN) and/or 3-deoxy-D-mannonic acid (DMN) represented by formula (VII) or the salts of these compounds



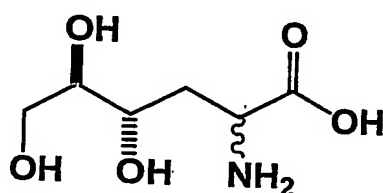
(VII)

The products resulting from this reaction undergo (oxidative) decarboxylation, preferably using hydrogen peroxide, to yield DRI. Production of a mixture of DGN and DMN or their salts takes place from KDG or its salts by reduction.

Preferably the preparation of 2-deoxy-D-ribose (DRI) is carried out by non-stereoselective reduction of 2-dehydro-3-deoxy-D-gluconic acid in water with sodium borohydride at room temperature using the method described for 2-keto-3-deoxyheptonic acid by Weissbach (J. Biol. Chem. 234 (1959), 705-709), followed by oxidative decarboxylation of 3-deoxy-D-gluconate and 3-deoxy-D-mannonate with hydrogen peroxide as described e.g. in US patent 3,312,683; Richards J. Chem. Soc. (1954), 3638-3640; Sowden J. Am. Chem. Soc. 76 (1954), 3541-3542.

In another preferred embodiment, the preparation of a mixture of DGN and DMN is carried out by hydrogenation of 2-dehydro-3-deoxy-D-gluconate in aqueous solution with 6% mol Nickel Raney catalyst or Platinum oxide at room temperature under a pressure of 6 bars.

In another embodiment of the invention, KDG or its salt is converted to 3-deoxy-D-glucosaminate (DGM) or 3-deoxy-D-mannosaminate (DMM) represented by formula (VIII) or the salts of these compounds



(VIII)

The products resulting from this reaction undergo (oxidative) decarboxylation, preferably using ninhydrin, to yield DRI. Production of a mixture of DGM and DMM or their salts takes place from KDG or its salts by reductive amination.

Preferably the preparation of 2-deoxy-D-ribose is carried out by non-stereoselective reductive amination of sodium or potassium 2-dehydro-3-deoxy-D-gluconate in

aqueous solution with ammonia and sodium cyanoborohydride at room temperature, followed by oxidative decarboxylation of 3-deoxy-D-2-glucosamine and 3-deoxy-D-2-mannosamine with ninhydrin using the method described for the synthesis of 2-deoxy-D-allose by Shelton (J. Am. Chem. Soc. 118 (1996), 2117-2125; and Borch, J. Am. Chem. Soc. 93 (1971), 2897; Durrwachter, J. Am. Chem. Soc. 108 (1986), 7812 referenced therein).

Furthermore, the present invention provides a method for producing the compound of formula (III) (2-deoxy-D-ribose) by converting the compound of formula (I) or one of its salts (KDG) in one single step. Preferably this conversion is achieved through enzymatic catalysis. This conversion is preferably catalysed by a keto acid decarboxylase. Preferred keto acid decarboxylases are thiamin pyrophosphate (TPP) dependent keto acid decarboxylases. Examples of TPP dependent keto acid decarboxylases are pyruvate decarboxylase (EC 4.1.1.1), a benzoylformate decarboxylase (EC 4.1.1.7), an indolepyruvate decarboxylase (EC 4.1.1.74), a phosphonopyruvate decarboxylase, a sulfopyruvate decarboxylase (EC 4.1.1.79), an oxalyl-coenzymeA decarboxylase (EC 4.1.1.8), an oxoglutarate decarboxylase (EC 4.1.1.71) or a phenylpyruvate decarboxylase (EC 4.1.1.43). It could be shown that keto acid decarboxylases, e.g., pyruvate decarboxylase enzymes from different organisms, can convert KDG into 2-deoxy-D-ribose (see Examples 8 to 12). In principle any keto acid decarboxylase can be used in connection with the present invention.

In a preferred embodiment of the method according to the invention KDG is converted into 2-deoxy-D-ribose by use of an enzyme having pyruvate decarboxylase activity.

A pyruvate decarboxylase catalyses the following reaction:



Several pyruvate decarboxylases (PDC) have been characterized as well as the corresponding *pdc* genes, for instance PDC from *Zymomonas mobilis* (Genbank accession number AAD19711; Neale et al., J. Bacteriol. 1987, 169:1024-1028), PDC from *Saccharomyces cerevisiae* (Genbank accession number NP013145; Candy et al., J. Gen. Microbiol. 1991, 137:2811-2815), PDC from *Acetobacter pasteurianus* (Genbank accession number AAM21208; Raj et al., Arch. Microbiol. 2001, 176:443-451), PDC from *Zymobacter palmae* (Genbank accession number AAM49566; Raj et al., Appl. Environ. Microbiol. 2002, 68:2869-2876), PDC from *Sarcina ventriculi* (Genbank accession number AAL18557; Lowe et al., J. Gen. Microbiol. 1992, 138:803-807). Many other pyruvate decarboxylases seems to occur in plants, fungi and bacteria as evidenced by the occurrence in these organisms of genes sharing sequence homologies with well-established *pdc* genes. Examples of such putative pyruvate decarboxylases are:

**PDC from plants:**

- Arabidopsis thaliana* (Genbank accession number T48155)
- Echinochloa crus-galli* (Genbank accession number AAM18119)
- Oryza sativa* (Genbank accession number NP922014)
- Rhizopus oryzae* (Genbank accession number AAM73540)
- Lotus corniculatus* (Genbank accession number AAO72533)
- Zea mays* (Genbank accession number BAA03354)
- Pisum sativum* (Genbank accession number CAA91445)
- Garden pea (Genbank accession number S65470)
- Nicotiana tabaccum* (Genbank accession number CAA57447)
- Solanum tuberosum* (Genbank accession number BAC23043)
- Fragaria ananassa* (Genbank accession number AAL37492)
- Cucumis melo* (Genbank accession number AAL33553)
- Vitis vinifera* (Genbank accession number AAG22488)

**PDC from Fungi :**

- Saccharum officinarum* (Genbank accession number CAB61763)
- Aspergillus oryzae* (Genbank accession number AAD16178)
- Aspergillus parasiticus* (Genbank accession number P51844)
- Saccharomyces cerevisiae* (Genbank accession number NP013145)
- Flammulina velutipes* (Genbank accession number AAR00231)

*Saccharomyces kluyveri* (Genbank accession number AAP75899)  
*Schizosaccharomyces pombe* (Genbank accession number CAB75873)  
*Candida glabrata* (Genbank accession number AAN77243)  
*Neurospora crassa* (Genbank accession number JN0782)  
*Pichia stipitis* (Genbank accession number AAC03164)  
*Kuyveromyces lactis* (Genbank accession number CAA61155)  
*Emericella nidulans* (Genbank accession number AAB63012)

**PDC from Prokaryotes:**

*Mycobacterium bovis* (Genbank accession number CAD93738)  
*Mycobacterium leprae* (Genbank accession number CAC31122)  
*Mycobacterium tuberculosis* (Genbank accession number NP215368)  
*Mycoplasma penetrans* (Genbank accession number NP758077)  
*Clostridium acetobutylicum* (Genbank accession number NP149189)  
*Acetobacter pasteurianus* (Genbank accession number AAM21208)  
*Zymobacter palmae* (Genbank accession number AAM49566)  
*Zymomonas mobilis* (Genbank accession number AAD19711)  
*Sarcina ventriculi* (Genbank accession number AAL18557)  
*Nostoc punctiforme* (Genbank accession number ZP00110850)

Such enzymes can be easily produced by recombinant microorganisms overexpressing the corresponding gene. Examples of genes coding for TPP dependent keto acid decarboxylases are *pdc* from *Zymomonas mobilis* (Genbank accession number AF124349), *pdc* from *Saccharomyces cerevisiae* (Genbank accession number NC001144), *pdc* from *Acetobacter pasteurianus* (Genbank accession number AF368435), *pdc* from *Zymobacter palmae* (Genbank accession number AF474145), *pdc* from *Sarcina ventriculi* (Genbank accession number AF354297). Other *pdc* genes can be found at Genbank corresponding to the above list of putative pyruvate decarboxylases.

In a preferred embodiment the pyruvate decarboxylase is of eukaryotic origin, more preferably it is from yeast and most preferably it is from *Saccharomyces cerevisiae*. In a particularly preferred embodiment the pyruvate decarboxylase is the pyruvate decarboxylase from *S. cerevisiae* which has the amino acid sequence as shown in SEQ ID NO: 21 (see also GenBank accession number NP013145).

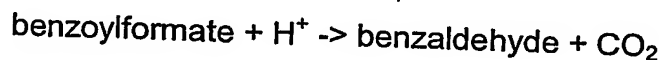
In another preferred embodiment the pyruvate decarboxylase is of prokaryotic origin, more preferably it is from an organism of the genus *Zymomonas* and most preferably from *Zymomonas mobilis*. In a particularly preferred embodiment the pyruvate decarboxylase is the pyruvate decarboxylase from *Z. mobilis* which has the amino acid sequence as shown in SEQ ID NO: 19 (see also GenBank accession number AAD19711).

In another preferred embodiment the prokaryotic pyruvate decarboxylase is from an organism of the genus *Acetobacter*, more preferably from the species *Acetobacter pasteurianus*. Particularly preferred the pyruvate decarboxylase is that of *A. pasteurianus* which shows the amino acid sequence as given in SEQ ID NO: 25 (see also GenBank accession number AAM21208).

In a further preferred embodiment the pyruvate decarboxylase is from an organism of the genus *Zymobacter*, more preferably of the species *Zymobacter palmae*. Particularly preferred is a pyruvate decarboxylase from *Z. palmae* which shows the amino acid sequence given in SEQ ID NO: 29 (see also GenBank accession number AAM49566).

In another preferred embodiment of the method according to the invention KDG is converted into 2-deoxy-D-ribose by use of an enzyme having benzoylformate decarboxylase activity.

A benzoylformate decarboxylase catalyses the following reaction:



A benzoylformate decarboxylase (BDC) from *Pseudomonas putida* (Genbank accessing number AAC15502; Tsou et al., Biochemistry. 1990, 29:9856-9862) has been characterized as well as the corresponding gene *mdlC* (Genbank accessing number AY143338). This enzyme has been shown to decarboxylate both D and L isomers of 2-keto-4,5-dihydroxyvalerate into the respective isomers of 3,4-

dihydroxybutanal (Niu et al., J. Am. Chem. Soc. 125 (2003), 12998-12999). Many other benzoylformate decarboxylases seems to occur in bacteria and archaeobacteria as evidenced by the occurrence in these organisms of genes sharing sequence homologies with genes coding for well-established BDC. Examples of such putative benzoylformate decarboxylases are:

**BDC from bacteria:**

*Pseudomonas aeruginosa* (Genbank accession number NP\_253588)  
*Rhodopseudomonas palustris* (Genbank accession number NP\_946955)  
*Streptomyces coelicolor* (Genbank accession number NP\_631486)  
*Chromobacterium violaceum* (Genbank accession number NP\_902771)  
*Bradyrhizobium japonicum* (Genbank accession number NP\_774243)

**BDC from archaeobacteria:**

*Sulfolobus solfataricus* (Genbank accession number NP\_343070)  
*Thermoplasma acidophilum* (Genbank accession number NP\_393976)  
*Thermoplasma volcanium* (Genbank accession number NP\_111716)

Such enzymes can be easily produced by recombinant microorganisms overexpressing the corresponding bdc gene. Such genes can be found at Genbank corresponding to the above list of putative benzoylformate decarboxylases.

Another example for a thiamine dependent decarboxylase which can be used in the method according to the invention is phosphonopyruvate decarboxylase. Several phosphonopyruvate decarboxylases (PPD) have been characterized as well as the corresponding genes, for instance PPD from *Bacteroides fragilis* (Genbank accession number AAG26466; Zhang et al., J. Biol. Chem. 2003, 278:41302-41308), PPD from *Streptomyces wedmorensis* (Genbank accession number BAA32496; Nakashita et al., J. Antibiot. 1997, 50:212-219). Many other phosphonopyruvate decarboxylases seem to occur in bacteria as evidenced by the occurrence in these organisms of genes sharing sequence homologies with genes coding for well-established PPD. Examples of such putative phosphonopyruvate decarboxylases are: PPD from *Bacteroides thetaiotaomicron* (Genbank accession number NP\_810632), PPD from *Amycolatopsis orientalis* (Genbank accession number CAB45023), PPD from *Clostridium tetani* E88 (Genbank accession number

NP\_782297), PPD from *Streptomyces viridochromogenes* (Genbank accession number CAA74722), PPD from *Streptomyces hygroscopicus* (Genbank accession number BAA07055), PPD from *Streptomyces coelicolor* A3 (Genbank accession number NP\_733715), *Streptomyces rishiriensis* (Genbank accession number AAG29796), *Bordetella pertussis* (Genbank accession number CAE 41214. Such enzymes can be easily produced by recombinant microorganisms overexpressing the corresponding gene.

A further example of a thiamine dependent decarboxylases which can be used in the method according to the present invention is sulfopyruvate decarboxylase. A sulfopyruvate decarboxylases (SPD) from *Methanococcus jannaschii* (Graupner et al., J. Bacteriol. 2000. 182:4862-4867) consisting of two subunits ComD (Genbank accession number P58415) and ComE (Genbank accession number P58416) has been characterized as well as the corresponding genes. Many other sulfopyruvate decarboxylases seems to occur in archaeobacteria and in bacteria as evidenced by the occurrence in these organisms of genes sharing sequence homologies with genes coding for well-established SPD.

Another further example of thiamine dependent decarboxylase which can be used in the method according to the present invention is indolepyruvate decarboxylase. Several indolepyruvate decarboxylases (IPD) have been characterized as well as the corresponding genes, for instance IPD from *Enterobacter cloacae* (Genbank accession number BAA14242; Scutz et al., 2003, Eur. J. Biochem. 270:2322-2331), IPD from *Azospirillum brasilense* (Genbank accession number AAC36886; Costacurta et al., Mol. Gen. Genet. 1994, 243:463-472), IPD from *Erwinia herbicola* (Genbank accession number AAB06571; Brandl et al., Appl. Environ. Microbiol. 1996, 62:4121-4128). Many other indolepyruvate decarboxylases seem to occur in bacteria as evidenced by the occurrence in these organisms of genes sharing sequence homologies with genes coding for well-established IPD.

Still another further example of a thiamine dependent decarboxylases which can be used in the method according to the present invention is phenylpyruvate decarboxylase. A phenylpyruvate decarboxylase from yeast (Genbank accession



number NP010668; Vuralhan et al., Appl. Environ. Microbiol. 2003, 69:4534-41) has been characterized as well as the corresponding gene ARO10 (Genbank accession number NC001136).

In a preferred embodiment of the method according to the invention in which the decarboxylation step is effected by an enzymatic reaction, the pH value is regulated by addition of an acid to be between pH 5 and pH 9, preferably between pH 6 and pH 8. In principle, any suitable acid can be used for this purpose. Preferred acids are HCl, H<sub>2</sub>SO<sub>4</sub>, D-gluconic acid or 2-dehydro-3-deoxy-D-gluconic acid.

Another aspect of the invention is a convenient and cost-effective method for preparing KDG either from D-gluconate (GCN) or from D-glucosamine through the use of recombinant enzymes.

In a preferred embodiment of the method of the invention, the compound of formula (I) is produced in a preliminary step from a D-gluconate salt by the use of a D-gluconate dehydratase activity. Preferred salts are potassium or sodium D-gluconate. Preferably the D-gluconate dehydratase is encoded by a polynucleotide comprising the nucleotide sequence selected from the group consisting of:

- (a) nucleotide sequences encoding a polypeptide comprising the amino acid sequence of SEQ ID N°2;
- (b) nucleotide sequences comprising the coding sequence of SEQ ID N°1;
- (c) nucleotide sequences encoding a fragment encoded by a nucleotide sequence of (a) or (b);
- (d) nucleotide sequences hybridising with a nucleotide sequence of any one of (a) to (c); and
- (e) nucleotide sequences which deviate from the nucleoside sequence of (d) as a result of degeneracy of the genetic code.

The enzymatic synthesis of KDG or its salts using D-gluconate dehydratase proceeds according to the following reaction: D-gluconate is converted into KDG by the elimination of one water molecule. The activity of a D-gluconate dehydratase has been characterized in different bacterial species e.g. in *Alcaligenes* (Kerstens,

Methods in Enzymology 42 (1975), 301-304); *Clostridium pasteurianum*, (Gottschalk, Methods in Enzymology 90 (1982), 283-287); *Thermoplasma acidophilum* (Budgen, FEBS Letters 196 (1986), 207-210) and *Sulfolobus solfataricus* (Nicolaus, Biotechnology Letters 8(7) (1986), 497-500). The preferred D-gluconate dehydratase was identified by screening several collection strains for D-gluconate dehydratase activity. The gene encoding a D-gluconate dehydratase, which was designated *gcnD* was selected from a genomic library of *Agrobacterium tumefaciens* strain C58, and further inserted in a multi copy vector optimised for expression. It was shown that a crude extract from *E. coli* cells over-expressing the *gcnD* gene catalysed the total conversion of D-gluconate into KDG (see Example 2).

In a further preferred embodiment of the method of the invention, the compound of formula (I) is produced in a preliminary step from D-glucosamine by the use of a D-glucosamine deaminase activity. Preferably the D-glucosamine deaminase is encoded by a polynucleotide comprising the nucleotide sequence selected from the group consisting of:

- (f) nucleotide sequences encoding a polypeptide comprising the amino acid sequence of SEQ ID N°4;
- (g) nucleotide sequences comprising the coding sequence of SEQ ID N°3;
- (h) nucleotide sequences encoding a fragment encoded by a nucleotide sequence of (a) or (b);
- (i) nucleotide sequences hybridising with a nucleotide sequence of any one of (a) to (c); and
- (j) nucleotide sequences which deviate from the nucleoside sequence of (d) as a result of degeneracy of the genetic code.

The enzymatic synthesis of KDG or its salts using D-glucosamine deaminase proceeds according to the following reaction: D-glucosamine is converted into KDG by the elimination of one molecule water and one molecule of ammonia. The activity of a D-glucosamine deaminase has been characterized in different bacterial species e.g. in *Pseudomonas fluorescens* (Iwamoto, Agric. Biol. Chem. 53 (1989), 2563-2569) *Agrobacterium radiobacter* (Iwamoto, FEBS Letters 104 (1979), 131-134;

Iwamoto, J. Biochem. 91 (1982), 283-289), and its requirement for  $Mn^{2+}$  ion was shown (Iwamoto, Biosci. Biotech. Biochem. 59 (1995), 408-411).

The preferred D-glucosamine deaminase was identified by screening several collection strains for D-glucosamine deaminase activity. The gene encoding a D-glucosamine deaminase, which was designated gmaA was isolated from *Agrobacterium tumefaciens* strain C58 by cloning a gene annotated as a putative D-serine deaminase. The gmaA gene was further inserted in a multi copy vector optimised for expression. It was shown that a crude extract from *E. coli* cells over-expressing the gmaA gene catalysed the conversion of D-glucosamine into KDG (see Example 4).

In a preferred embodiment the present invention relates to a method for producing a compound of formula III, in particular 2-deoxy-D-ribose, starting from D-gluconate or D-glucosamine by enzymatic reactions which, in a first step, convert D-gluconate or D-glucosamine into KDG as described above and, in a second step, convert KDG into 2-deoxy-D-ribose as described above.

Thus, the enzymatic conversion of D-gluconate into KDG can be achieved by the use of a D-gluconate dehydratase. The enzymatic conversion of D-glucosamine into KDG can be achieved by the use of a D-glucosamine deaminase. With respect to the preferred embodiments the same applies as has already been set forth above.

The enzymatic conversion of the resulting KDG into 2-deoxy-D-ribose can be achieved by the use of a keto acid decarboxylase. With respect to the preferred embodiments the same applies as has been set forth above.

The enzymatic two step method of converting D-gluconate or D-glucosamine into 2-deoxy-D-ribose via KDG can be carried out in vitro by using cell extracts of cells expressing the corresponding enzymes or by using purified or partially purified enzymes. The enzymes can be enzymes which are naturally expressed in an organism or they may be recombinantly produced. Methods of preparing and isolating corresponding (recombinant) enzymes are well-known to the person skilled in the art.

In a preferred embodiment the enzymatic two step method of converting D-gluconate or D-glucosamine into 2-deoxy-D-ribose via KDG is carried out in vivo, i.e. by using a suitable organism, which expresses the required enzyme activities. This organism

may be any type of organism, preferably it is a cell, e.g. a plant, an animal, a fungal cell or a bacterial cell. Most preferably fungal or bacterial cells are used. Preferred fungi are yeasts, such as *Saccharomyces cerevisiae*; preferred bacterial cells are, e.g. *E. coli*, *Zymomonas mobilis*, *Zymobacter palmae*, *Acetobacter pasteurianus*, *Acinetobacter calcoaceticus*, *Agrobacterium tumefaciens* and *Bacillus subtilis*. The organism may be an organism which endogenously already expresses one of the enzymatic activities, i.e. a D-gluconate dehydratase or a D-glucosamine deaminase for producing KDG, or a keto acid decarboxylase for converting KDG into 2-deoxy-D-ribose, and in which the respective other enzymatic activity is expressed due to the introduction of a corresponding exogenous nucleic acid molecule encoding the corresponding enzyme. Alternatively, the organism may also be an organism which naturally does not express the enzyme activities required for converting D-gluconate or D-glucosamine into KDG and further into 2-deoxy-D-ribose and in which corresponding foreign nucleic acid molecules have been introduced encoding D-gluconate dehydratase or D-glucosamine deaminase and a keto acid decarboxylase, respectively.

In a particularly preferred embodiment the organism is an organism which does not express a KDG kinase (kdgK) activity. Such an enzyme activity would lead to a phosphorylation of KDG to KDPG, which in turn is cleaved by an aldolase into pyruvate and glyceraldehyde-phosphate, thereby diverting KDG into a different unwanted metabolic pathway. It is possible to use for the method according to the invention organisms which naturally do not express a kdgK gene. If the used organism naturally expresses a kdgK, means and methods are well-known to the skilled person to produce mutants or variants of such an organism in which the corresponding kdgK gene is inactivated.

If the described method according to the invention is carried out in vivo by using an organism which expresses a D-gluconate dehydratase for converting D-gluconate into KDG and a keto acid decarboxylase for converting KDG into 2-deoxy-D-ribose, this has the advantage that one can provide D-gluconate as a substrate in the culture medium used to culture the organism. D-gluconate is taken up by the organism and is converted into 2-deoxy-D-ribose.

In another particularly, preferred embodiment the organism is an organism which does not express a KDG aldolase (encoded by the *eda* gene in *E. coli*) activity. Such an enzyme activity would lead to cleavage of KDG into pyruvate and glyceraldehydes, thereby diverting KDG into a different unwanted metabolic pathway. It is possible to use for the method according to the invention organisms which naturally do not express an *eda* gene. If the used organism expresses an *eda* gene, means and methods are well-known to the skilled person to produce mutants or variants of such an organism in which the corresponding *eda* gene is inactivated.

In still another particularly, preferred embodiment the organism is an organism which does not express a 2-deoxy-D-ribose aldolase (encoded by the *deoC* gene in *E. coli*) activity. Such an enzyme activity would lead to cleavage of 2-deoxy-D-ribose into acetaldehyde and glyceraldehyde, thereby diverting 2-deoxy-D-ribose into a different unwanted metabolic pathway. It is possible to use for the method according to the invention organisms which naturally do not express a *deoC* gene. If the used organism expresses a *deoC* gene, means and methods are well-known to the skilled person to produce mutants or variants of such an organism in which the corresponding *deoC* gene is inactivated. For instance a *deoC* mutant of *E. coli* has been reported (Valentin-Hansen, EMBO J. 1 (1982), 317-322) as well as a method for deleting the *deo* operon in *E. coli* (Kaminski, J. Biol. Chem. 277 (2002), 14400-14407; Valentin-Hansen, Molec. Gen. Genet. 159 (1978), 191-202).

The present invention also relates to organism which are capable of enzymatically converting D-gluconate into KDG due to the expression of a D-gluconate dehydratase and/or of enzymatically converting D-glucosamine into KDG due to the expression of a D-glucosamine deaminase and which are furthermore capable of enzymatically converting KDG into 2-deoxy-D-ribose by a decarboxylation reaction catalysed by a keto acid decarboxylase. The organism may in principle be any suitable organism, preferably, it is a cell, e.g. a plant cell, an animal cell, a fungal cell or a bacterial cell. More preferably, it is a fungal or a bacterial cell. Preferred fungi are yeasts, e.g. *Saccharomyces cerevisiae*. Preferred bacteria are *Escherichia coli*, *Zymomonas mobilis*, *Zymobacter palmae*, *Acetobacter pasteurianus*, *Acinetobacter calcoaceticus*, *Agrobacterium tumefaciens* and *Bacillus subtilis*. In one aspect, the

organism is an organism which already endogenously expresses a D-gluconate dehydratase or a D-glucosamine deaminase and into which a foreign nucleic acid molecule has been introduced which encodes a keto acid decarboxylase which can catalyse the decarboxylation of KDG to 2-deoxy-D-ribose. With respect to the preferred embodiments of the keto acid decarboxylase the same applies as has been set forth previously.

In another aspect, the organism is an organism which already expresses a keto acid decarboxylase which is capable of converting KDG into 2-deoxy-D-ribose by a decarboxylation reaction but which does not naturally express a D-gluconate dehydratase or a D-glucosamine deaminase, and into which a foreign nucleic acid molecule has been introduced which encodes a D-gluconate dehydratase and/or which encodes a D-glucosamine deaminase. I.e. the organism can be genetically modified so as to express a D-gluconate dehydratase or a D-glucosamine deaminase or both enzymes.

In a further aspect, the organism is an organism, which naturally does not express a D-gluconate dehydratase, a D-glucosamine deaminase and a keto acid decarboxylase which is capable of converting KDG by decarboxylation into 2-deoxy-D-ribose, and into which foreign nucleic acid molecules have been introduced encoding a D-gluconate dehydratase or a D-glucosamine deaminase, or both, and a nucleic acid molecule which encodes a keto acid decarboxylase which is capable of converting KDG into 2-deoxy-D-ribose by decarboxylation.

With respect to the preferred embodiments of the D-gluconate dehydratase, the D-glucosamine deaminase and the keto acid decarboxylase to be expressed in the organisms according to the invention, the same applies which has been set forth above in connection with the method according to the invention.

In a particularly preferred embodiment the organism according to the invention does not express a KDG kinase (kdgK) activity. It can either be an organism which naturally does not express kdgK or it can be an organism which naturally expresses a kdgK but in which the corresponding gene has been inactivated, e.g. by gene disruption or other suitable methods well-known to the person skilled in the art.

The present invention also relates to the use of an enzyme having keto acid decarboxylase activity or of a polynucleotide encoding such an enzyme in a method for converting KDG into 2-deoxy-D-ribose. With respect to the preferred embodiments the same applies as has already been set forth in connection with the method according to the present invention.

These and other embodiments are disclosed and encompassed by the description and examples of the present invention. The disclosure content of any references cited above or below is herewith incorporated into the present application. Further literature concerning any one of the methods, uses and compounds to be employed in accordance with the present invention may be retrieved from public libraries, using for example electronic devices. For example the public database "Medline" may be utilized which is available on the Internet, for example under <http://www.ncbi.nlm.nih.gov/PubMed/medline.html>. Further databases and addresses, such as <http://www.ncbi.nlm.nih.gov/>, <http://www.infobiogen.fr/>, [http://www.fmi.ch/biology/research\\_tools.html](http://www.fmi.ch/biology/research_tools.html), <http://www.tigr.org/>, are known to the person skilled in the art and can also be obtained using, e.g., <http://www.google.de>. An overview of patent information in biotechnology and a survey of relevant sources of patent information useful for retrospective searching and for current awareness is given in Berks, TIBTECH 12 (1994), 352-364.

Furthermore, the term "and/or" when occurring herein includes the meaning of "and", "or" and "all or any other combination of the elements connected by said term".

## EXAMPLES

### **Example 1: Cloning of a gene encoding a D-gluconate dehydratase from *Agrobacterium tumefaciens* strain C58 (CIP 104333)**

*Agrobacterium tumefaciens* strain C58 (CIP 104333) was obtained from Institut Pasteur Collection (CIP, Paris, France). Chromosomal DNA was extracted and a D-gluconate dehydratase gene was amplified by PCR according to standard protocols using the following primers:

5'-CCCTTAATTAATGACGACATCTGATAATCTTC-3', depicted in SEQ ID N° 5;

5'-TTTGCGGCCGCTTAGTGGTTATCGCGCGGC-3', depicted in SEQ ID N° 6;

5'-CCCGGTACCATGACGACATCTGATAATCTTC-3', depicted in SEQ ID N° 7;

A first DNA fragment amplified using the two primers depicted in SEQ ID N° 5 and SEQ ID N° 6, was ligated into a pUC18-derived vector previously digested by *PacI* and *NotI* to yield the plasmid pVDM80. A second DNA fragment amplified using the two primers depicted in SEQ ID N° 6 and SEQ ID N° 7, was ligated into a pET29a vector (Novagen) previously digested by *KpnI* and *NotI* to yield the plasmid pVDM82. The nucleotide sequence of the cloned gene is depicted in SEQ ID N° 1 and the sequence of the polypeptide encoded by this gene is depicted in SEQ ID N° 2.

**Example 2: Expression of a D-gluconate dehydratase activity in *Escherichia coli* and preparation of 2-dehydro-3-deoxy-D-gluconate from D-gluconate**

Competent cells of *E. coli* BL21 were transformed with the pVDM82 plasmid constructed as described in example 1 yielding strain +1289. Strain + 1289 was cultivated at 30°C in Luria-Bertani (LB) medium (Difco) containing 30 mg/l kanamycin until OD(600 nm) reached a value of 0.6. Then isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) was added to a 0.5 mM final concentration. After a further cultivation period of 2 hours and 30 minutes, cells were collected by centrifugation and washed once with 20 mM sodium phosphate buffer pH 7.2. A cell extract was prepared by suspending about 5 g of cells in 10 ml of Tris-HCl 50 mM pH 8.5 buffer containing 10000 units lysozyme (Ready-Lyse, Epicentre, Madison, Wisconsin) and 1 mM EDTA, and incubating the suspension at 30°C for 15 minutes. Then 10000 kUnits deoxyribonuclease I (DNase I, Sigma) as well as 5 mM MgCl<sub>2</sub> were added to the preparation which was incubated at 30°C for an additional period of 15 minutes. The cell extract thus obtained was kept frozen at -20°C before use.

1.5 ml of the cell extract was mixed with 2M sodium or potassium D-gluconate in a total volume of 10 ml. This preparation was incubated at 37°C after the pH has been adjusted to 8.5. The progression of 2-dehydro-3-deoxy-D-gluconate (KDG) synthesis was followed by analysing aliquots taken after increasing periods of incubation. Several dilution parts of these aliquots were deposited on silica plates and chromatographed in the following solvent system: isopropanol / water (90/10). A yellow spot of KDG (R<sub>f</sub> ~0.40) was detected after revelation with p-anisaldehyde.



KDG was also quantitated using a spectrophotometric assay based on the reaction with semicarbazide hydrochloride as described by Mac Gee (J. Biol. Chem. 1954. 210, 617-626). Typically, after a 30h period of incubation and using the spectrophotometric assay, KDG concentration ranged from 1.5 to 2 M.

The sodium or potassium 2-dehydro-3-deoxy-D-gluconate solution thus obtained could be used as such for further synthetic steps. 2-Dehydro-3-deoxy-D-gluconic acid could also be prepared from such a solution applying published protocols (Bender, Anal. Biochem. 1974. 61, 275-279). A crude preparation of a mixture of 2-dehydro-3-deoxy-D-gluconic acid and KCl could also be obtained by adding one equivalent of HCl to a potassium 2-dehydro-3-deoxy-D-gluconate solution which was then evaporated.

**Example 3: Cloning of a gene encoding a D-glucosaminase from *Agrobacterium tumefaciens* strain C58 (CIP 104333)**

*Agrobacterium tumefaciens* strain C58 (CIP 104333) was obtained from Institut Pasteur Collection (CIP, Paris, France). Chromosomal DNA was extracted and a D-glucosaminase gene was amplified by PCR according to standard protocols using the following primers:

5'-CCCTTAATTAATGCAGTCTTCTTCAGCTCTTC-3', depicted in SEQ ID N° 8;  
5'-TTTGCGGCCGCTAGTGAAAGAAGGTTGTGTAGAT-3', depicted in SEQ ID N° 9;

5'-AAATCATGACTATGCAGTCTTCTTCAGCTCTTCG-3', depicted in SEQ ID N° 10;  
5'-TATAGATCTCTAGTGAAAGAAGGTTGTGTAGAT-3', depicted in SEQ ID N° 11;

A first DNA fragment amplified using the two primers depicted in SEQ ID N° 8 and SEQ ID N° 9, was ligated into a pUC18-derived vector previously digested by *PacI* and *NotI* to yield the plasmid pKDGb1. A second DNA fragment amplified using the two primers depicted in SEQ ID N° 10 and SEQ ID N° 11, was ligated into a pQE60 vector (Qiagen) previously digested by *BspH1* and *BglII* to yield the plasmid pEP18. The nucleotide sequence of the cloned gene is depicted in SEQ ID N° 3 and the sequence of the polypeptide encoded by this gene is depicted in SEQ ID N° 4.

**Example 4: Expression of a D-glucosaminase deaminase activity in Escherichia coli and preparation of 2-dehydro-3-deoxy-D-gluconic acid from D-glucosamine**

Competent cells of E. coli MG1655 were transformed with the pEP18 plasmid constructed as described in example 1 and pREP4 (Qiagen) yielding strain +1068. Strain + 1068 was cultivated at 37°C in LB medium containing 30 mg/l kanamycin and 100 mg/l ampicillin until OD(600 nm) reached a value of 0.6. Then IPTG was added to a 0.5 mM final concentration. After a further cultivation period of 2 hours and 30 minutes, cells were collected by centrifugation and washed once with 20 mM sodium phosphate buffer pH 7.2. A cell extract was prepared using the protocol described in example 2.

2 ml of the cell extract was mixed with 100 mM sodium or potassium D-glucosamine and 0.1 mM pyridoxal phosphate in a total volume of 5 ml. This preparation was incubated at 37°C after the pH has been adjusted to 7.5.

The progression of 2-dehydro-3-deoxy-D-gluconate (KDG) synthesis was followed using the protocols described in example 2. Typically, after a 30h period of incubation and using the spectrophotometric assay described in example 2, KDG concentration ranged from 50 to 100 mM.

**Example 5: Preparation of 2-deoxy-D-ribonate from 2-dehydro-3-deoxy-D-gluconate**

0.5 ml of a 31% hydrogen peroxide solution were added to 5 ml of a 1M potassium 2-dehydro-3-deoxy-D-gluconate (KDG) solution at 25°C. The progression of KDG decarboxylation was followed both by the observation of bubbles resulting from the release of carbon dioxide and by the disappearance of KDG using the thin layer chromatography protocol described in example 2. Typically, after a 3h period of reaction the concentration of residual KDG was less than 10 mM.

**Example 6: Preparation of 2-deoxy-D-ribitol from 2-deoxy-D-ribonolactone**

0.2 g of Rhodium ( 5 % on carbon) catalyst was added to an aqueous solution of 1 g 2-deoxy-D-ribonolactone prepared following a method described by Deriaz (J. Chem. Soc. (1949), 1879-1883) for the synthesis of 2-deoxy-L-ribonolactone. Hydrogenation of 2-deoxy-D-ribonolactone was performed at 130°C under a pressure of 80 bars. The solution obtained after filtration of the reaction mixture was evaporated. The residue was dissolved in ethyl acetate and further purified by chromatography on a silica column. The solvent was removed in vacuo leading to a yellow oil (yield 85%). The compound thus obtained was identical with 2-deoxy-D-ribitol obtained by reduction of 2-deoxy-D-ribose as described by Rabow (J. Am. Chem. Soc. 122 (1999), 3196-3203).

**Example 7: Preparation of 1-N-morpholino-3,4,5-trihydroxypentene-1 from 2-dehydro-3-deoxy-D-gluconate**

2 g of 2-dehydro-3-deoxy-D-gluconic acid were suspended in 150 ml benzene. 1.1 ml morpholine and 100 mg p-toluenesulfonic acid were added to the suspension and the reaction mixture was refluxed for 3 hours. Water formed by this reaction was removed by distillation. Benzene was decanted. Solid compounds attached to the vessel were collected, washed with acetone and dried. The main compound present in this preparation (yield 40%) was further purified by column chromatography on a silica column using a gradient of methanol in chloroform. Fractions containing 1-N-morpholino-3,4,5-trihydroxypentene-1 were pooled and solvent was removed in vacuo.

<sup>1</sup>H-NMR (D<sub>2</sub>O): δ = 3.15 ppm (4H, t, morpholine), 3.8 ppm (4H, t, morpholine), 3.4 to 4 ppm, (4H, m, 5a-H, 5b-H, 4-H, 3-H), 6.3 and 6.8 ppm (2H, 2d, 1-H and 2-H, J = 4 Hz).

**Example 8: Cloning of a gene encoding a pyruvate decarboxylase from *Zymomonas mobilis***

*Zymomonas mobilis* strain B-806 (CIP 102538T) was obtained from Institut Pasteur Collection (CIP, Paris, France). Chromosomal DNA was extracted and a pyruvate

decarboxylase gene was amplified by PCR according to standard protocols using the following primers:

5'-GCGTTAATTAATGAGTTATACTGTCGGTACC-3', depicted in SEQ ID N° 12;

5'-TATGCGGCCGCTTAGAGGAGCTTGTTAACAGG-3', depicted in SEQ ID N° 13;

The DNA fragment amplified using the two primers depicted in SEQ ID N° 12 and SEQ ID N° 13, was ligated either into pSP100 or into pEVL5 (respectively a pUC18-derived or a pQE70-derived vector as described below) previously digested by *PacI* and *NotI* to yield respectively plasmid pEVL107 and plasmid pEVL420. The nucleotide sequence of the cloned gene as well as the encoded sequence of the corresponding polypeptide can be found at GenBank (accession number AF124349) and is shown in SEQ ID NO: 18 and SEQ ID NO: 19, respectively.

Plasmid pSP100 was obtained by introducing a ribosomal binding site, a *PacI* and a *NotI* restriction sites into a pUC18 vector previously digested by *EcoRI* and *BamHI* using standard protocols. The complete nucleotide sequence of pSP100 is depicted in SEQ ID N° 14.

Plasmid pEVL5 was obtained by introducing a ribosomal binding site, a *PacI* and a *NotI* restriction sites into a pQE70 vector (Qiagen) previously digested by *EcoRI* and *BamHI* using standard protocols. The complete nucleotide sequence of pEVL5 is depicted in SEQ ID N° 15.

**Example 9: Cloning of a gene encoding a pyruvate decarboxylase from *Saccharomyces cerevisiae***

Chromosomal DNA was extracted from *Saccharomyces cerevisiae* strain S288C (ATCC 204508) and a pyruvate decarboxylase gene was amplified by PCR according to standard protocols using the following primers:

5'-ATATTTAATTAATGTCTGAAATTACTTTGG-3', depicted in SEQ ID N° 16;

5'-ATATGCGGCCGCTTATTGCTTAGCGTTGGT-3', depicted in SEQ ID N° 17;

The DNA fragment amplified using the two primers depicted in SEQ ID N° 16 and SEQ ID N° 17, was ligated either into pSP100 or into pEVL5 (respectively a pUC18-derived or a pQE70-derived vector as described in example 8) previously digested by

PacI and NotI to yield respectively plasmid pVDM61 and plasmid pEVL419. The nucleotide sequence of the cloned gene as well as the encoded sequence of the corresponding polypeptide can be found at GenBank (accession number NC001144) and is shown in SEQ ID NO: 20 and SEQ ID NO: 21, respectively.

**Example 10: Expression of a pyruvate decarboxylase activity in Escherichia coli and enzymatic synthesis of 2-deoxy-D-ribose from 2-dehydro-3-deoxy-D-gluconate**

**Expression of pyruvate decarboxylase and preparation of cell-free extracts**

Competent cells of E. coli MG1655 strain were transformed with either pEVL107 or pVDM61 (constructed as described in examples 8 and 9) yielding respectively strain +1735 and strain + 844. These strains were cultivated at 37°C in Luria-Bertani (LB) medium (Difco) containing 100 mg/l ampicillin until OD(600 nm) reached a value around 1.5.

Competent cells of E. coli MG1655 strain harbouring pREP4 plasmid (Qiagen) were transformed with either pEVL420 or pEVL419 (constructed as described in Examples 8 and 9) yielding respectively strain +3150 and +3148. These strains were cultivated at 37°C in Luria-Bertani (LB) medium (Difco) containing 100 mg/l ampicillin and 30 mg/l kanamycin until OD(600 nm) reached a value of 0.6. Then isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a 0.5 mM final concentration. After a further cultivation period of 2 hours and 30 minutes, cells were collected by centrifugation and washed once with 20 mM sodium phosphate buffer pH 7.2.

For each strain a cell-free extract was prepared using the same protocol as described in Example 2. Then crude cell-free extracts were passed through a PD-10 column (Amersham) equilibrated with 50 mM Tris-acetate buffer pH 6 and stored at -20°C.

**Enzymatic synthesis of 2-deoxy-D-ribose from 2-dehydro-3-deoxy-D-gluconate**

1.0 ml of cell-free extract was mixed with 20 mM sodium 2-dehydro-3-deoxy-D-gluconate, 0.5 mM thiamine pyrophosphate and 5 mM MgCl<sub>2</sub> in a total volume of 1.5 ml of 50 mM tris-acetate buffer pH 6. The progression of 2-deoxy-D-ribose (DRI) synthesis was followed by analysing aliquots taken after increasing periods of

Incubation at 37°C. About 1 µl of each aliquot which had been previously concentrated five-fold by evaporation was deposited on a silica plate and chromatographed in the following solvent system: butanol / triethylamine / water (10/2/5). A blue spot of DRI (Rf ~0.50) was detected after revelation with orcinol when using cell-free extracts of either strain +3150 or +3148 after a period of incubation of 65 hours. The crude preparation containing the spot corresponding to DRI was concentrated and passed through a 1.5 ml silica column equilibrated with isopropanol. The fractions containing the expected DRI compound were pooled, concentrated and the resulting sample analysed by mass spectrometry. The results of such an analysis confirmed the identity of the isolated compound with DRI, and the production of DRI from KDG catalysed by pyruvate decarboxylase either from *Zymomonas mobilis* or from *Saccharomyces cerevisiae*.

**Example 11: Cloning of a gene encoding a pyruvate decarboxylase from *Acetobacter pasteurianus*, expression of encoded pyruvate decarboxylase activity in *Escherichia coli* and enzymatic synthesis of 2-deoxy-D-ribose from 2-dehydro-3-deoxy-D-gluconate**

*Acetobacter pasteurianus* strain NCIB 8618 (DSMZ 2347) was obtained from DSMZ Collection (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany). Chromosomal DNA was extracted from the cells and a pyruvate decarboxylase gene was amplified by PCR according to standard protocols using the following primers:

5'-TCTTTAATTAATGGGTTGTCCGTCATTCATATA-3', depicted in SEQ ID N° 22;

5'-CTAAAGCTTTTAGGCCAGAGTGGTCTTGCGCG-3', depicted in SEQ ID N° 23;

The DNA fragment amplified using the two primers depicted in SEQ ID N° 22 and SEQ ID N° 23, was ligated either into pSP100 or into pEVL5 (respectively a pUC18-derived or a pQE70-derived vector as described in example 8) previously digested by *PacI* and *NotI* to yield respectively plasmid pEVL541 and plasmid pEVL560. The nucleotide sequence SEQ ID N° 24 of the cloned gene as well as the encoded sequence of the corresponding polypeptide SEQ ID N° 25 can be found at GenBank (accession number AF368435).

Competent cells of *E. coli* MG1655 strain were transformed with pEVL541 yielding strain +3559. Competent cells of *E. coli* MG1655 strain harbouring pREP4 plasmid (Qiagen) were transformed with pEVL560 yielding strain +3924. These strains were cultivated and cell-free extracts were prepared as described in Example 10. Cell-free extracts were incubated with KDG and the progression of 2-deoxy-D-ribose (DRI) synthesis was followed as described in Example 10. A spot corresponding to DRI was observed indicating that pyruvate decarboxylase from *Acetobacter pasteurianus* was able to decarboxylate KDG into DRI.

**Example 12: Cloning of a gene encoding a pyruvate decarboxylase from *Zymobacter palmae*, expression of encoded pyruvate decarboxylase activity in *Escherichia coli* and enzymatic synthesis of 2-deoxy-D-ribose from 2-dehydro-3-deoxy-D-gluconate**

*Zymobacter palmae* strain T109 (DSMZ10491) was obtained from DSMZ Collection (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany). Chromosomal DNA was extracted from the cells and a pyruvate decarboxylase gene was amplified by PCR according to standard protocols using the following primers:

5'-ATCTTAATTAATGTATACCGTTGGTATGTACT-3', depicted in SEQ ID N° 26;  
5'-TATGCGGCCGCTTACGCTTGTGGTTTGCGAGAGT-3', depicted in SEQ ID N° 27;

The DNA fragment amplified using the two primers depicted in SEQ ID N° 26 and SEQ ID N° 27, was ligated either into pSP100 or into pEVL5 (respectively a pUC18-derived or a pQE70-derived vector as described in example 8) previously digested by *PacI* and *NotI* to yield respectively plasmid pEVL546 and plasmid pEVL561. The nucleotide sequence of the cloned gene as well as the encoded sequence of the corresponding polypeptide is shown in SEQ ID NOs: 28 and 29, respectively and can be found at GenBank (accession number AF474145).

Competent cells of *E. coli* MG1655 strain were transformed with pEVL546 yielding strain +3568. Competent cells of *E. coli* MG1655 strain harbouring pREP4 plasmid (Qiagen) were transformed with pEVL560 yielding strain +3923. These strains were cultivated and cell-free extracts were prepared as described in Example 10. Cell-free

extracts were incubated with KDG and the progression of 2-deoxy-D-ribose (DRI) synthesis was followed as described in Example 10. A spot corresponding to DRI was observed indicating that pyruvate decarboxylase from *Zymobacter palmae* was able to decarboxylate KDG into DRI.

**Example 13. Cloning of a gene encoding a benzoylformate decarboxylase from *Pseudomonas putida*, expression of encoded benzoylformate decarboxylase activity in *Escherichia coli* and enzymatic synthesis of 2-deoxy-D-ribose from 2-dehydro-3-deoxy-D-gluconate.**

*Pseudomonas putida* strain Migula (DSMZ 291) was obtained from DSMZ Collection (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany). Chromosomal DNA was extracted from and a benzoylformate decarboxylase gene was amplified by PCR according to standard protocols using the following primers:

5'-CTATTAATTAATGGCTTCGGTACACGGCACCA-3', depicted in SEQ ID N° 30;  
5'-TATGCGGCCGCTTACTTCACCGGGCTTACGGTGC-3', depicted in SEQ ID N° 31;

The DNA fragment amplified using the two primers depicted in SEQ ID N° 30 and SEQ ID N° 31, was ligated either into pSP100 or into pEVL5 (respectively a pUC18-derived or a pQE70-derived vector as described in example 8) previously digested by *PacI* and *NotI* to yield respectively plasmid pEVL681 and plasmid pEVL670. The nucleotide sequence SEQ ID N° 32 of the cloned gene as well as the encoded sequence of the corresponding polypeptide SEQ ID N° 33 can be found at GenBank (accession number AY143338).

Competent cells of *E. coli* MG1655 strain were transformed with pEVL681 yielding strain +4050. Competent cells of *E. coli* MG1655 strain harbouring pREP4 plasmid (Qiagen) were transformed with pEVL670 yielding strain +3927. Those strains were cultivated and cell-free extracts were prepared as described in example 10. Cell-free extracts were incubated with KDG and the progression of 2-deoxy-D-ribose (DRI) synthesis was followed as described in example 10. A spot corresponding to DRI was observed indicating that benzoylformate decarboxylase from *Pseudomonas putida* was able to decarboxylate KDG into DRI.



#### Preparative enzymatic synthesis of 2-deoxy-D-ribose

100 µl of cell-free extract from strain +3927 (containing 2.5 mg of bacterial proteins) were mixed with 300 mM sodium 2-dehydro-3-deoxy-D-gluconate, 0.5 mM thiamine pyrophosphate and 5 mM MgCl<sub>2</sub> in a total volume of 0.5 ml of 80 mM potassium phosphate buffer pH 6. After a period of incubation of 16 and 40 hours, few µl of a solution of HCl 2N were added to the incubation mixture until the pH reached a value of 6. The progression of 2-deoxy-D-ribose (DRI) synthesis was also followed by analysing aliquots taken after increasing periods of incubation at 37°C. About 1 µl of each aliquot was deposited on a silica plate and chromatographed as described in example 10. The concentration of 2-deoxy-D-ribose was estimated to be about 200 mM by comparison with standard solutions. <sup>13</sup>C NMR analysis of the crude mixture confirmed that the compound formed from 2-dehydro-3-deoxy-D-gluconate was 2-deoxy-D-ribose, and that the concentration of 2-deoxy-D-ribose was closed to 25 g/l. Another preparative enzymatic synthesis was performed in the same conditions except that no addition of acid was made along the incubation period. In those conditions, the concentration of 2-deoxy-D-ribose was closed to 10 g/l, far lower than the concentration reached in the preceding experiment for which the pH had been controlled and regularly adjusted to a value of 6.

#### **Example 14. Enzymatic synthesis of 2-deoxy-D-ribose from D-gluconate**

One pot enzymatic synthesis of 2-deoxy-D-ribose from D-gluconate was achieved as follows, using D-gluconate dehydratase encoded by gcnD gene of *Agrobacterium tumefaciens* and pyruvate decarboxylase from *Zymomonas mobilis*:

50 µl of cell-free extract from strain +1289 (containing 1.5 mg of bacterial proteins) and 400 µl of cell-free extract from strain +3150 (containing 17 mg of bacterial proteins after concentration by ultrafiltration) prepared as described respectively in example 2 and in example 10, were mixed with 50 mM potassium D-gluconate, 0.5 mM thiamine pyrophosphate and 5 mM MgCl<sub>2</sub> in a total volume of 0.5 ml of 50 mM N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid) (HEPES) buffer pH 7. The progression of 2-deoxy-D-ribose (DRI) synthesis was also followed by analysing

aliquots taken after increasing periods of incubation at 37°C. After a period of incubation of 18 hours, about 1 µl of the incubation mixture was deposited on a silica plate and chromatographed as described in example 10. The concentration of 2-deoxy-D-ribose was estimated to be about 1 g/l by comparison with standard solutions.